

0550
1108

#4

OIKE

RAW SEQUENCE LISTING

DATE: 11/21/2001

PATENT APPLICATION: US/09/818,066

TIME: 12:09:10

Input Set : N:\Crf3\RULE60\09818066.txt

Output Set: N:\CRF3\11212001\I818066.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

- 6 (i) APPLICANT: Shuping Tong et al.
- 8 (ii) TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
- 10 (iii) NUMBER OF SEQUENCES: 75
- 12 (iv) CORRESPONDENCE ADDRESS:
 - 13 (A) ADDRESSEE: Fish & Richardson P.C.
 - 14 (B) STREET: 225 Franklin Street
 - 15 (C) CITY: Boston
 - 16 (D) STATE: Massachusetts
 - 17 (E) COUNTRY: U.S.A.
 - 18 (F) ZIP: 02110-2804

ENTERED

20 (v) COMPUTER READABLE FORM:

- 21 (A) MEDIUM TYPE: Floppy disk
- 22 (B) COMPUTER: IBM PC compatible
- 23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- 24 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

26 (vi) CURRENT APPLICATION DATA:

- C--> 27 (A) APPLICATION NUMBER: US/09/818,066
- C--> 28 (B) FILING DATE: 27-Mar-2001
- 34 (C) CLASSIFICATION:

31 (vii) PRIOR APPLICATION DATA:

- 32 (A) APPLICATION NUMBER: US 08/683,262
- 33 (B) FILING DATE: 18-JUL-1996

36 (viii) ATTORNEY/AGENT INFORMATION:

- 37 (A) NAME: Fraser, Janis K.
- 38 (B) REGISTRATION NUMBER: 31,819
- 39 (C) REFERENCE/DOCKET NUMBER: 00786/287002

41 (ix) TELECOMMUNICATION INFORMATION:

- 42 (A) TELEPHONE: (617) 542-5070
- 43 (B) TELEFAX: (617) 542-8906
- 44 (C) TELEX: 200154

46 (2) INFORMATION FOR SEQ ID NO: 1:

48 (i) SEQUENCE CHARACTERISTICS:

- 49 (A) LENGTH: 27 base pairs
- 50 (B) TYPE: nucleic acid
- 51 (C) STRANDEDNESS: single
- 52 (D) TOPOLOGY: linear

54 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

56 GCAGATCTAT GGGCAGAATC TTTCCAC

27

58 (2) INFORMATION FOR SEQ ID NO: 2:

60 (i) SEQUENCE CHARACTERISTICS:

- 61 (A) LENGTH: 25 base pairs
- 62 (B) TYPE: nucleic acid
- 63 (C) STRANDEDNESS: single
- 64 (D) TOPOLOGY: linear

66 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

RAW SEQUENCE LISTING

DATE: 11/21/2001

PATENT APPLICATION: US/09/818,066

TIME: 12:09:10

Input Set : N:\Crf3\RULE60\09818066.txt

Output Set: N:\CRF3\11212001\I818066.raw

```

68 GTGAATTCAG CGCAGGGTCC CCAAT                                     25
71 (2) INFORMATION FOR SEQ ID NO: 3:
73     (i) SEQUENCE CHARACTERISTICS:
74         (A) LENGTH: 28 base pairs
75         (B) TYPE: nucleic acid
76         (C) STRANDEDNESS: single
77         (D) TOPOLOGY: linear
79     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
81 TCAGATCTAT GATGGGGCAA CATCCAGC                                     28
83 (2) INFORMATION FOR SEQ ID NO: 4:
85     (i) SEQUENCE CHARACTERISTICS:
86         (A) LENGTH: 30 base pairs
87         (B) TYPE: nucleic acid
88         (C) STRANDEDNESS: single
89         (D) TOPOLOGY: linear
91     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
93 GCGAATTCAG GTACCAGACA TTTTCTTCTT                                   30
95 (2) INFORMATION FOR SEQ ID NO: 5:
97     (i) SEQUENCE CHARACTERISTICS:
98         (A) LENGTH: 27 base pairs
99         (B) TYPE: nucleic acid
100        (C) STRANDEDNESS: single
101        (D) TOPOLOGY: linear
103     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
105 GCGAATTCTT ATTCCTAACT CTTGTAA                                     27
107 (2) INFORMATION FOR SEQ ID NO: 6:
109     (i) SEQUENCE CHARACTERISTICS:
110        (A) LENGTH: 20 base pairs
111        (B) TYPE: nucleic acid
112        (C) STRANDEDNESS: single
113        (D) TOPOLOGY: linear
115     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
117 GARYTNTAYG TNATGGAGAT                                             20
119 (2) INFORMATION FOR SEQ ID NO: 7:
121     (i) SEQUENCE CHARACTERISTICS:
122        (A) LENGTH: 23 base pairs
123        (B) TYPE: nucleic acid
124        (C) STRANDEDNESS: single
125        (D) TOPOLOGY: linear
127     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
129 AAYTCNGGYT CNCCNGCYTC RTG                                         23
131 (2) INFORMATION FOR SEQ ID NO: 8:
133     (i) SEQUENCE CHARACTERISTICS:
134        (A) LENGTH: 21 base pairs
135        (B) TYPE: nucleic acid
136        (C) STRANDEDNESS: single
137        (D) TOPOLOGY: linear
139     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
141 TKYTNAGYCA YGARTTYCAR G                                           21

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/818,066

DATE: 11/21/2001

TIME: 12:09:10

Input Set : N:\Crf3\RULE60\09818066.txt

Output Set: N:\CRF3\11212001\I818066.raw

```

143 (2) INFORMATION FOR SEQ ID NO: 9:
145     (i) SEQUENCE CHARACTERISTICS:
146         (A) LENGTH: 20 base pairs
147         (B) TYPE: nucleic acid
148         (C) STRANDEDNESS: single
149         (D) TOPOLOGY: linear
151     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
153 TTKGCNGART ANARNGTYTC                                     20
155 (2) INFORMATION FOR SEQ ID NO: 10:
157     (i) SEQUENCE CHARACTERISTICS:
158         (A) LENGTH: 20 base pairs
159         (B) TYPE: nucleic acid
160         (C) STRANDEDNESS: single
161         (D) TOPOLOGY: linear
163     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
165 ATGAAACAGA CACTGAAGAA                                     20
167 (2) INFORMATION FOR SEQ ID NO: 11:
169     (i) SEQUENCE CHARACTERISTICS:
170         (A) LENGTH: 20 base pairs
171         (B) TYPE: nucleic acid
172         (C) STRANDEDNESS: single
173         (D) TOPOLOGY: linear
175     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
177 ATGGAGATCT CGGACGGCCC                                     20
179 (2) INFORMATION FOR SEQ ID NO: 12:
181     (i) SEQUENCE CHARACTERISTICS:
182         (A) LENGTH: 20 base pairs
183         (B) TYPE: nucleic acid
184         (C) STRANDEDNESS: single
185         (D) TOPOLOGY: linear
187     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
189 TTCTTCAGTG TCTGTTTCAT                                     20
192 (2) INFORMATION FOR SEQ ID NO: 13:
194     (i) SEQUENCE CHARACTERISTICS:
195         (A) LENGTH: 9 amino acids
196         (B) TYPE: amino acid
197         (D) TOPOLOGY: linear
199     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
201 Gly Ser Arg Arg Ala Ser Val Gly Ser
202 1           5
204 (2) INFORMATION FOR SEQ ID NO: 14:
206     (i) SEQUENCE CHARACTERISTICS:
207         (A) LENGTH: 30 base pairs
208         (B) TYPE: nucleic acid
209         (C) STRANDEDNESS: single
210         (D) TOPOLOGY: linear
212     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
214 ATCACTGAGC TCAAATTACC CCATGAGATG                         30
216 (2) INFORMATION FOR SEQ ID NO: 15:

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/818,066

DATE: 11/21/2001

TIME: 12:09:10

Input Set : N:\Crf3\RULE60\09818066.txt

Output Set: N:\CRF3\11212001\I818066.raw

```

218      (i) SEQUENCE CHARACTERISTICS:
219          (A) LENGTH: 30 base pairs
220          (B) TYPE: nucleic acid
221          (C) STRANDEDNESS: single
222          (D) TOPOLOGY: linear
224      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:
226 GGAAACTCGA GCTGGAAGCA GTGTATGAA                               30
228 (2) INFORMATION FOR SEQ ID NO: 16:
230      (i) SEQUENCE CHARACTERISTICS:
231          (A) LENGTH: 33 base pairs
232          (B) TYPE: nucleic acid
233          (C) STRANDEDNESS: single
234          (D) TOPOLOGY: linear
236      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:
238 ATGGTACCAT GGAGGCGGCG CGGTGCATCG AGC                               33
241 (2) INFORMATION FOR SEQ ID NO: 17:
243      (i) SEQUENCE CHARACTERISTICS:
244          (A) LENGTH: 31 base pairs
245          (B) TYPE: nucleic acid
246          (C) STRANDEDNESS: single
247          (D) TOPOLOGY: linear
249      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:
251 ATCTCGAGAT ATTAACATTA GCAATGTTAC T                               31
254 (2) INFORMATION FOR SEQ ID NO: 18:
256      (i) SEQUENCE CHARACTERISTICS:
257          (A) LENGTH: 16 amino acids
258          (B) TYPE: amino acid
259          (D) TOPOLOGY: linear
261      (ii) MOLECULE TYPE: protein
263      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:
265 Gln Trp Thr Pro Glu Asp Gln Lys Ala Arg Glu Ala Phe Arg Arg
266  1      5              10              15
268 (2) INFORMATION FOR SEQ ID NO: 19:
270      (i) SEQUENCE CHARACTERISTICS:
271          (A) LENGTH: 27 amino acids
272          (B) TYPE: amino acid
273          (D) TOPOLOGY: linear
275      (ii) MOLECULE TYPE: protein
277      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:
279 Ser Val Glu Leu Arg Glu Leu Tyr Val Met Glu Ile Ser Asp Asn Pro
280  1      5              10              15
282 Gly Val His Glu Ala Gly Glu Pro Glu Phe Lys
283      20              25
285 (2) INFORMATION FOR SEQ ID NO: 20:
287      (i) SEQUENCE CHARACTERISTICS:
288          (A) LENGTH: 20 amino acids
289          (B) TYPE: amino acid
290          (D) TOPOLOGY: linear
292      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/818,066

DATE: 11/21/2001

TIME: 12:09:10

Input Set : N:\Crf3\RULE60\09818066.txt

Output Set: N:\CRF3\11212001\I818066.raw

```

294 Leu Ile Asp Arg Thr Arg Ile Val Ile Val Pro Ser Leu Asn Pro Asp
295 1 5 10 15
297 Gly Arg Ile Ala
298 20
300 (2) INFORMATION FOR SEQ ID NO: 21:
302 (i) SEQUENCE CHARACTERISTICS:
303 (A) LENGTH: 22 amino acids
304 (B) TYPE: amino acid
305 (D) TOPOLOGY: linear
307 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:
309 Ser Leu Leu Ser His Glu Phe Gln Asp Glu Thr Asp Thr Glu Glu Glu
310 1 5 10 15
312 Thr Leu Tyr Ser Ala Lys
313 20
315 (2) INFORMATION FOR SEQ ID NO: 22:
317 (i) SEQUENCE CHARACTERISTICS:
318 (A) LENGTH: 13 amino acids
319 (B) TYPE: amino acid
320 (D) TOPOLOGY: linear
322 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:
324 Val Glu Glu Gly Lys Val Pro Val Leu Asn Thr Pro Asp
325 1 5 10
327 (2) INFORMATION FOR SEQ ID NO: 23:
329 (i) SEQUENCE CHARACTERISTICS:
330 (A) LENGTH: 22 amino acids
331 (B) TYPE: amino acid
332 (D) TOPOLOGY: linear
334 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:
336 Glu Leu Tyr Val Met Glu Ile Ser Asp Asn Pro Gly Val His Glu Ala
337 1 5 10 15
339 Gly Glu Pro Glu Phe Lys
340 20
342 (2) INFORMATION FOR SEQ ID NO: 24:
344 (i) SEQUENCE CHARACTERISTICS:
345 (A) LENGTH: 66 base pairs
346 (B) TYPE: nucleic acid
347 (C) STRANDEDNESS: single
348 (D) TOPOLOGY: linear
350 (ix) FEATURE:
355 (A) NAME/KEY: misc_feature
352 (B) LOCATION: 3..3
353 (D) OTHER INFORMATION: /note= "R is A or G."
375 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:
377 GARCTNTAYG TNATGGARAT WAGYGAYAAY CCNGGNGTNC AYGARGCNGG NGARCCNGAR 60
379 TTAAAR 66
381 (2) INFORMATION FOR SEQ ID NO: 25:
383 (i) SEQUENCE CHARACTERISTICS:
384 (A) LENGTH: 66 base pairs
385 (B) TYPE: nucleic acid

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/818,066

DATE: 11/21/2001

TIME: 12:09:11

Input Set : N:\Crf3\RULE60\09818066.txt

Output Set: N:\CRF3\11212001\I818066.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:1982 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:1986 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:1990 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:1994 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:1998 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2002 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2006 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2010 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2014 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2018 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2022 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2026 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2030 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2034 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2038 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2042 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2046 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2050 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2054 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2058 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2062 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2066 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2070 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2074 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:63
L:2175 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2179 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2183 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2187 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2191 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2195 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2199 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2203 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2207 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2211 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:65
L:2369 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=74